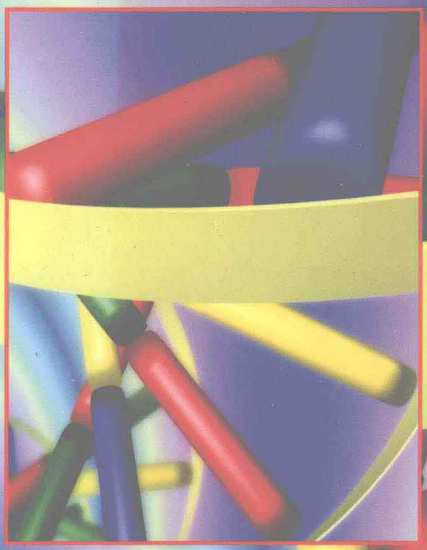
Fourth Edition

# Understanding DNA and Gene Cloning

## A Guide for the Curious



Karl Drlica

# Understanding DNA and Gene Cloning

## A GUIDE FOR THE CURIOUS

*Fourth Edition*

**KARL DRLICA**

*Public Health Research Institute*
*225 Warren Street*
*Newark, NJ 07103*

John Wiley & Sons, Inc.

| | |
|---|---|
| *Senior Acquisitions Editor* | Keri Witman |
| *Marketing Manager* | Clay Stone |
| *Production Manager* | Pamela Kennedy |
| *Production Editor* | Sarah Wolfman-Robichaud |
| *Senior Designer* | Kevin Murphy |
| *Illustration* | Matrix Art Services |
| *Cover Photo* | © Dennis Meyler/Corbis Stock Market |

This book was set in 10/12 Palatino by Matrix Publishing Services and printed and bound by Courier Westford. The cover was printed by Phoenix Color Corp.

This book is printed on acid-free paper.

The paper on this book was manufactured by a mill whose forest management programs include sustained yield harvesting of its timberlands. Sustained yield harvesting principles ensure that the number of trees cut each year does not exceed the amount of new growth.

The following material was adapted from *Double-Edged Sword: The Promises and Risks of the Genetic Revolution,* © 1994 by Karl A. Drlica and reprinted by permission of Addison-Wesley Longman Publishing Co., Inc.: descriptions of patterns of inheritance, Mendelian inheritance, DNA fingerprinting, and Figures 13-2, 13-3, 13-4, 14-3, and 14-5.



ISBN: 0471-43416-7

Drlica, Karl.
Understanding DNA and gene cloning : a guide for the curious / Karl Drlica.—4th ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-471-43416-7 (pbk.)
1. Molecular cloning. 2. Recombinant DNA. 3. Genetic engineering. 4. Genetics. I. Title.

QH442.2.D75 2003
576.5—dc21

2003043077

*To Ilene, for many years of support*

# PREFACE TO THE FOURTH EDITION

An explosion of knowledge has been shaking the science of biology for more than a decade. The explosion is now beginning to touch many of us in a very personal way, most often through DNA testing. At the center of the explosion is chemical information—information that our cells use, store, and pass on to subsequent generations. With our new understanding of life comes the ability to manipulate chemical information—we can restructure the molecules that program living cells. This new technology is being used to solve problems in such diverse areas as waste disposal, synthesis of drugs, treatment of cancer, plant breeding, and diagnosis of human diseases. The new biology is also telling us how the chemicals in our bodies function; we may soon be writing our own biological future. When this happens, each of us will be confronted with a new set of personal and political choices. Some of these difficult and controversial decisions are already upon us, and the choices will not get easier. Informed decisions require an understanding of molecular biology and recombinant DNA technology; this book is intended to provide that understanding.

Molecular biology is a science of complex ideas supported by test tube experiments with molecules. Consequently, the science has remained largely inaccessible to those without a knowledge of chemistry. I have tried to change that situation—this book requires the reader to have little or no background in chemistry. Consequently, our discussion of DNA starts at a much more elementary level than is commonly found in publications such as *Scientific American* and *The New York Times*. Chemical processes and molecular structures are described by means of analogies using terms familiar to nonsci-

entists. Technical terms are kept to a minimum; where they must be introduced, they are accompanied by definitions. These definitions are grouped, often in expanded form, in the glossary. To provide a feeling for informational molecules, I have also introduced a few details about how they are manipulated experimentally. Integration of these details should help remove the mystery from gene cloning and expose the elegance and simplicity of the technology.

This is the fouth edition of *Understanding DNA.* When I completed the first edition, gene cloning was largely an academic subject. By the time the second edition was finished, information about gene cloning had become commercially important: stock brokers, patent attorneys, and judges were making decisions that required knowledge of DNA and its activities. The first edition had even been cited as a reference for DNA patent cases by the U.S. Court of Appeals. With the third edition another quantum leap in molecular genetics had occurred, this time into the realm of health prediction. Everyone in the industrialized world needed to understand the implications of gene cloning to take advantage of the new technology and to minimize the risks associated with release of personal genetic information. Thus *Understanding DNA* was broadened to be useful for personal decision making. That meant adding chapters dealing with human genetics and restructuring the presentation. Now another quantum leap has occurred, this time allowing us to work with whole genomes (a genome is the entire information set of an organism). Although genomics is a field in its infancy, the associated technology is readily accessible to so many biologists that it is already rivaling genetics and biochemistry as a way to study life.

Originally, *Understanding DNA* was written for college students, with readers ranging from nonscience majors to potential biology majors. While the audience now extends to the general public, several elements have been retained to keep *Understanding DNA* appropriate for classroom use. Among these are "Questions for Discussion" at the end of each chapter. Some of these questions have specific answers to reinforce a point; others are open-ended to stimulate additional reading. Many of the questions introduce information that would dilute the main themes if included in the primary text. Another teaching aid is the glossary. Vocabulary is a key aspect of learning molecular biology, and the reader should expect to refer frequently to the definitions listed. As an instructor I found that ad-

ministering simple glossary quizzes early in a course overcomes some of the vocabulary barriers. A third aid is the list of additional readings, which has also been expanded. Many of the entries are from *Scientific American* because the articles are of high quality, they are at the appropriate level when used in conjunction with *Understanding DNA,* and *Scientific American* is readily available.

I thank the following persons for helping with the fourth edition: Shirley Chapin, Marila Gennaro, Wes Hatfield, Samuel Kayman, Tao Lu, Richard Pine, Richard Thaw, Peter Tolias, Ilene Wagner, and Xilin Zhao. I also thank the staff at John Wiley & Sons, including Catherine Donovan, Keri Witman, Deborah Herbert, Sarah Wolfman-Robichaud and Brenda Griffing for encouragement and skillful production of the work.

Karl Drlica

# INTRODUCTION

In the mid-1980s media coverage of the O.J. Simpson murder trial made DNA a household word, but few people fully appreciated how important DNA science would be to their own lives. One way to judge the importance is to notice how often DNA-based themes appear in movies. A hint of the growing importance can also be gleaned from casual reading of major newspapers. A typical medical example involves a baby who would have been plagued with infections due to a faulty immune system. The condition had been discovered by prenatal DNA testing, however, and when the baby was born, doctors saved cells from the umbilical cord. The cells were then engineered with DNA to correct the immune system defect. When placed back in the baby, those cells may allow the child to live a normal life. In another example, some of the families that participated in identifying a "breast cancer gene" were outraged at geneticists for withholding information about the breast cancer status of subjects' children. The scientists argued that a child is incapable of giving the informed consent that physicians require before they reveal such devastating news. The parents, however, wanted the information to help guide and hopefully protect their children.

A different type of example concerns personal identification. DNA analyses have entered the military through a program in which blood and tissue samples are routinely obtained from recruits for identification purposes in case war-related activities result in dismemberment. Fearing an invasion of genetic privacy, two marines chose court martial over participation in the program. In a different twist on authority, attorneys of men wrongly convicted of rape have had DNA analyses performed on the semen-stained clothing that had been used as evidence. In dozens of cases it was demonstrated that the semen in question could not have come from the defendant;

thus innocent men were released from prision. A British approach to tracking down rapists has involved thousands of men. Sometimes all the males in a community are asked to volunteer for DNA tests to flush out the perpetrator. The key point for each of these examples is that information in DNA can be used to track people and their diseases.

Before long, the medical profession will be able to look into our health future through a wide variety of genetic tests. Individuals will be identified who are especially susceptible to particular ailments, including problems arising from specific environmental dangers. A person who happens to be in a susceptible group will be advised to arrange his or her lifestyle to avoid the particular danger. For example, people who get skin cancer quickly when exposed to sunlight can reduce their risk by shading themselves and by staying indoors much of the time. Those who have a high probability of contracting breast or colon cancer can monitor their bodies more carefully than most people. Sometimes they opt for preemptive surgery. Still others are in danger of lethal reactions to anesthetics commonly used during surgery. The foreknowledge can help them avoid the danger. Even scientists are taking advantage of their own work. A woman working on a gene involved with heart disease tested her own blood and found that she was predisposed to the disease. She immediately started taking the vitamin that helps reduce the problem. The list of predictable, treatable problems will soon grow to where most of us can benefit from genetic testing.

Testing also has a downside: many of the maladies that can be identified currently have no cure. This poses a dilemma because people will not necessarily want to know that they are destined to suffer from an incurable disease decades from now. Doctors will not be able to tell them what to do. Even when a physician's advice can help us make a clear-cut decision, such assistance may be difficult to obtain because so many tests will be available. For example, proper counseling requires about 30 minutes per disease; for a hundred diseases we would need 3000 minutes. It will not be possible for each of us to obtain 50 hours of technical consultation from our doctors—we will be forced to use our own understanding of DNA and genetics to make decisions.

Another problem is genetic discrimination. In the past, families carrying genetic ailments such as Huntington disease were consid-

ered to be cursed. The uncontrollable shaking made it easy to identify the afflicted, who were sometimes incarcerated or put to death as witches. In modern times, it's usually careers and financial status that suffer as members of Huntington disease families are deemed uninsurable or are rejected from lengthy professional training programs. In the near future, improvements in the detection of many different genetic blemishes will reveal that each of us is predisposed to some condition we would prefer to avoid. As this knowledge reaches our employers and insurance carriers, they may be tempted to reduce their costs by cutting us from their rolls. Even when discrimination is illegal, we will have to learn to recognize the bias and take appropriate action.

Managed health care brings an additional dimension to the discrimination dilemma: for some of us, our doctors are employed by our insurance companies or by our employers. Press reports suggest that doctors are sometimes pressured by companies to reveal medical records of employees thereby feeding worries about breaches in doctor-patient confidentiality. The message is quite explicit: for some types of test, such as that for HIV infection and those for genetic disease, avoid your company doctor.

Just the fear of discrimination will make us cautious about what is written in our medical records, since we know that this material is not secure (life insurance companies often require us to release medical histories prior to issuing coverage). The fear will not be limited to ourselves, since genetic diseases are passed from one generation to the next. Thus we will worry that our disclosure will lead to problems for our children and grandchildren. And if we don't tell our doctors everything about our family medical history, we take on the burden of self-diagnosis. We will each have to learn medical genetics.

If we look a bit farther into the future, we see that it will be possible to change DNA information in individuals, in families, and in nations. At that point we will be confronted by a new set of decisions. At the individual level, gene therapy protocols are already adding good genes to body cells that have a bad one. Should this expensive treatment be made available to everyone through government subsidies? At the family level, genetic tracking can follow hereditary diseases as they pass from one generation to the next, enabling geneticists to single out persons harboring genetic disease

and assist them with family planning. Afflicted fetuses can often be identified early in pregnancy. However, the definition of "afflicted" is controversial, and there is disagreement over when it should be applied in individual abortion decisions. Since selective abortion could in principle eradicate a disease from a family, governments may be tempted to initiate genetic improvement programs as they seek ways to reduce health care costs. Here the decisions become quite profound because we as a society would be taking a small step toward altering the very nature of humankind.

As the Age of the Gene Hunter comes into full bloom, each of us will have the opportunity to get much more out of our bodies. At the same time, we will face a variety of risks, stemming largely from the ability of other people to know about our prospects for health. Difficult decisions lie ahead, decisions that require an understanding of DNA.

K. D.

# CONTENTS

CHAPTER ONE

# PREVIEW

## Life as Interacting Molecules

### *Overview*

Information governing the characteristics of all organisms is stored in long, thin molecules of deoxyribonucleic acid (DNA). DNA molecules contain regions (genes) that specify the structure of other molecules called proteins. Protein molecules in turn control cellular chemistry and contribute to cell structure. Biologists can obtain large amounts of specific regions of DNA. The general strategy involves first cutting the DNA into small fragments. The fragments are then moved into single-celled organisms, often bacteria or yeast. Conditions are set up so the fragments become a permanent part of the microorganism, and pure cultures of bacteria or yeast containing a particular DNA fragment are then obtained. The fragments are removed and used for further study, for detecting disease, for changing the cellular chemistry of another organism, and for producing large quantities of specific proteins of medical or industrial value. Cloning and nucleotide sequencing methods have also been used to determine the entire information set for a variety of organisms. This genomic information makes it possible to study the complex and subtle ways that genes interact with each other. Since microorganisms receiving DNA fragments are being changed in unknown ways, gene cloning experiments were originally viewed as *potentially* dangerous. Regulations were soon established to minimize the potential

health hazards. Many types of recombinant DNA have now been constructed and studied; no harmful effects have been observed, and almost all biologists now consider the cloning of genes to be safe. Use of the cloned genes, however, raises a variety of new issues that range from the uncontrolled spread of genetically engineered plants to genetic modification of humans.

## INTRODUCTION

In a general sense biologists have solved the riddle of **heredity**, the question of how offspring come to resemble their parents. The answer lies in the chemical behavior of submicroscopic structures called molecules. (All boldface words are defined in the glossary; italic is used for emphasis, for names of genes, and for scientific names of organisms.) At the center of this new understanding is a giant molecule called **deoxyribonucleic acid (DNA)**. This book is about DNA, the chemical that specifies features such as eye color and blood type. DNA in our cells influences all our physical characteristics, as is true for every living **organism** on earth. This book is also about **genetic engineering**, **recombinant DNA**, and **gene cloning**. In particular, it is about how gene cloning works and what we have learned from doing it. This first chapter initiates a discussion of DNA structure and function, and a few comments are made concerning fundamental features of **atoms**, **molecules**, **enzymes**, and **cells**. Gene cloning is introduced to provide a context for the discussion. The chapter concludes by outlining the types of information presented in the rest of the book.

The basic unit of life is the cell, an organized set of chemical reactions bounded by a **membrane** and capable of self-perpetuation. Our bodies are collections of trillions of cells that work together. For example, liver cells cluster to form livers, kidney cells form kidneys, and skin cells attach to each other to cover our bodies. With few exceptions, every cell contains all the information required for an independent existence; indeed, under the right conditions human cells can be removed from the body and grown in laboratory dishes. The information necessary to control the chemistry of the cell (i.e., the

chemistry of life) is stored in the long, thin fiber called DNA. DNA fibers are found in every cell of our bodies except mature red blood cells.

Isolated DNA looks like a tangled mass of string (Figure 1-1). Our cells, which are generally less than a **millimeter** long, contain about 2 meters of DNA specially packaged to fit inside. DNA can be bent, wrapped, looped, twisted, and even tied in knots. Many DNA molecules are circles that are sometimes found interlinked like a magician's rings. Thus DNA is very flexible, at least in terms of three-dimensional structure. But in terms of information content, DNA is quite rigid, for the same information passes from generation to generation with little change.

One of the goals of this book is to explain how information is stored in DNA, how it is reproduced, and how it is used. For now, the important concept is that distinct regions of DNA contain distinct bits of information. The specific regions of information are called **genes**. In some ways DNA is similar to motion picture film. Like film, DNA is subdivided into *frames* that make sense when seen in the correct order. In DNA the *frames* correspond to the letters in the genetic code, which is described in Chapter 2. When a number of frames or genetic letters are organized into a specific combination, they create a scene in the case of film and a gene in the case of DNA (Figure 1-2).

Information in genes is used primarily for the manufacture of **proteins**, chainlike molecules that fold in a precise way to form specific structures. Some proteins contribute to the architecture of the cell, while others directly control cell chemistry. Occasionally we can easily see the effects of particular genes and proteins. For example, a small group of genes carries the information for making the proteins that determine eye color. It is the specific information in the DNA, in the genes, that makes human beings different from honey bees and fir trees.

## ATOMS AND MOLECULES

Each DNA fiber is a molecule, a group of atoms joined together to form a distinct unit. Several points are important for understanding discussions of atoms and molecules. First, all forms of matter are

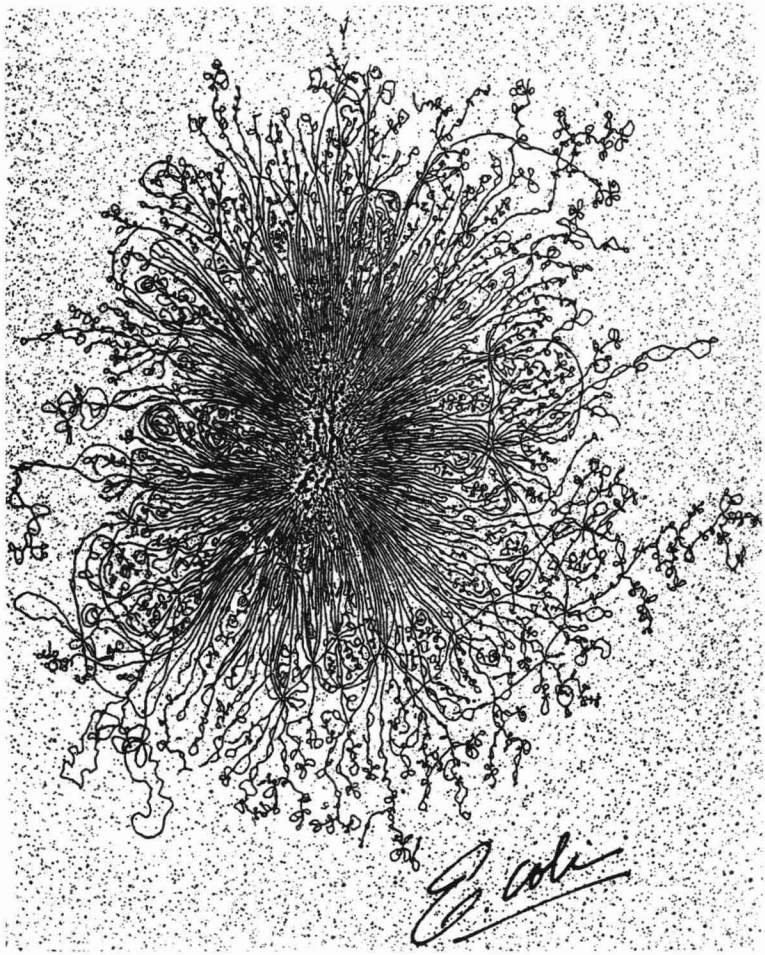



**Figure 1-1. Electron Micrograph of a DNA Molecule Released from a Bacterium.** The long, threadlike material is DNA, which is about one millimeter long or about 1000 times the length of the bacterium from which it was taken. The molecular details of how this DNA is compacted and packaged to fit in the cell are not yet understood. The electron micrograph is of a purified, surface-spread *E. coli* chromosome prepared by Ruth Kavenoff and Brian Bowen. The line under the *E. coli* signature represents 2.5 micrometers. (Copyright 1983 with all rights reserved by DesignerGenes in trust for the Julius Marmur Memorial Fund, Biochemistry Dept., Albert Einstein College of Medicine, Bronx, NY 10461, USA).